



US007616631B2

(12) **United States Patent**  
**Hetrick et al.**

(10) **Patent No.:** **US 7,616,631 B2**  
(45) **Date of Patent:** **Nov. 10, 2009**

(54) **METHOD AND APPARATUS FOR  
DEBUGGING PROTOCOL TRAFFIC  
BETWEEN DEVICES IN INTEGRATED  
SUBSYSTEMS**

(75) Inventors: **William A. Hetrick**, Wichita, KS (US);  
**Jeremy Dean Stover**, Wichita, KS (US);  
**Matt Tiemeyer**, Bothell, WA (US)

(73) Assignee: **LSI Corporation**, Milpitas, CA (US)

(\*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 1050 days.

6,188,686	B1 *	2/2001	Smith	370/388
6,343,324	B1 *	1/2002	Hubis et al.	709/229
6,400,730	B1 *	6/2002	Latif et al.	370/466
6,766,466	B1 *	7/2004	Jibbe	714/42
6,785,742	B1 *	8/2004	Teow et al.	710/1
6,834,311	B2 *	12/2004	Rao	370/352
7,159,216	B2 *	1/2007	McDonald	718/100
7,216,258	B2 *	5/2007	Ebsen et al.	714/30
7,533,288	B2 *	5/2009	Hatasaki et al.	714/4
2002/0064181	A1 *	5/2002	Ofek et al.	370/477
2002/0103913	A1 *	8/2002	Tawil et al.	709/229
2002/0170004	A1 *	11/2002	Parrett et al.	714/43
2003/0074599	A1 *	4/2003	Golasky et al.	714/6
2003/0131182	A1 *	7/2003	Kumar et al.	711/5

(21) Appl. No.: **10/218,726**

(22) Filed: **Aug. 14, 2002**

(65) **Prior Publication Data**

US 2004/0032834 A1 Feb. 19, 2004

(51) **Int. Cl.**  
**H04L 12/50** (2006.01)

(52) **U.S. Cl.** ..... **370/388**; 370/477; 370/906;  
714/33

(58) **Field of Classification Search** ..... 370/352,  
370/378-379, 388, 428, 457, 477, 906; 714/33  
See application file for complete search history.

(56) **References Cited**

**U.S. PATENT DOCUMENTS**

5,633,999 A \* 5/1997 Clowes et al. .... 714/6

\* cited by examiner

*Primary Examiner*—Ricky Ngo

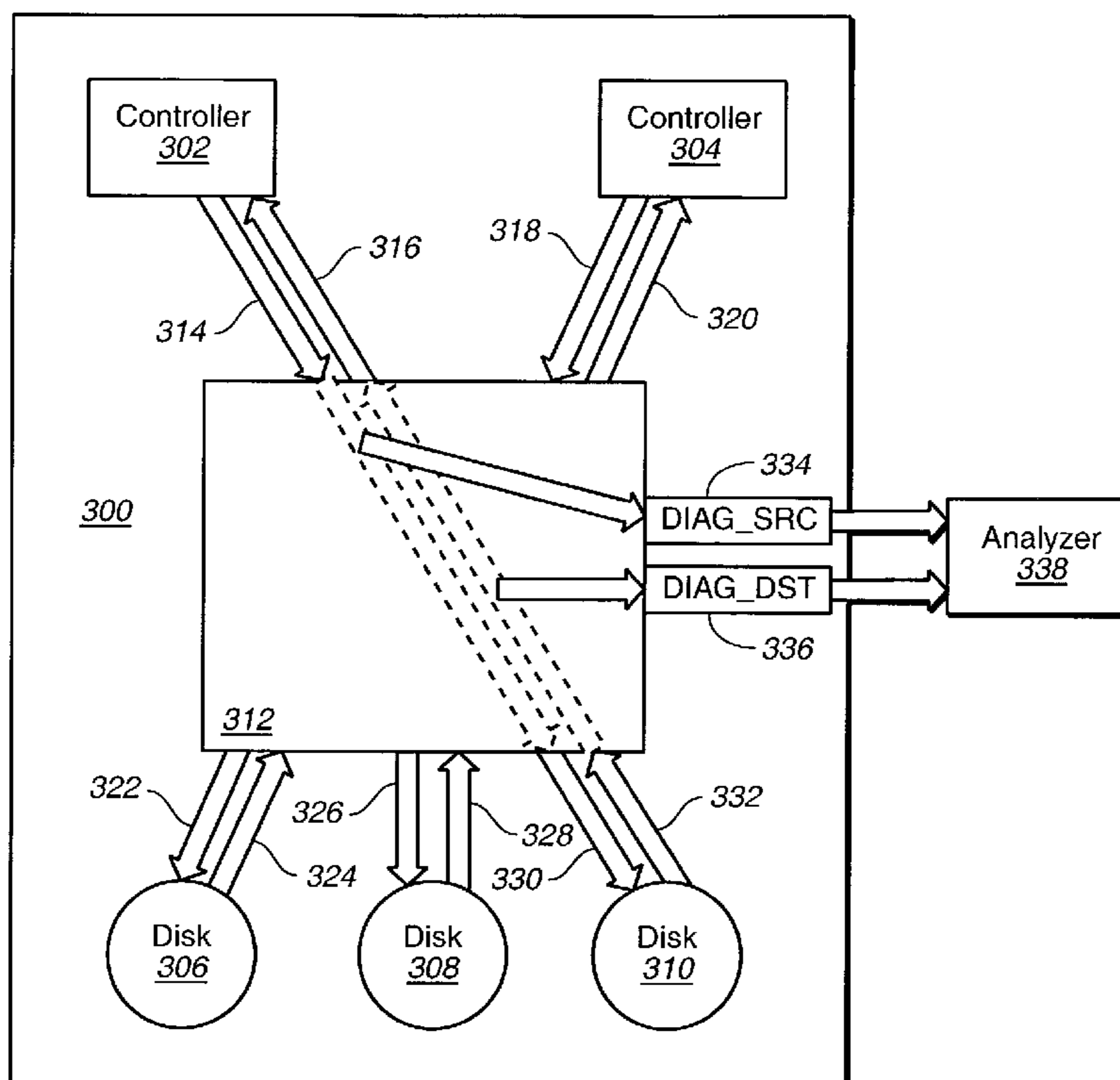
*Assistant Examiner*—Phuongchau B Nguyen

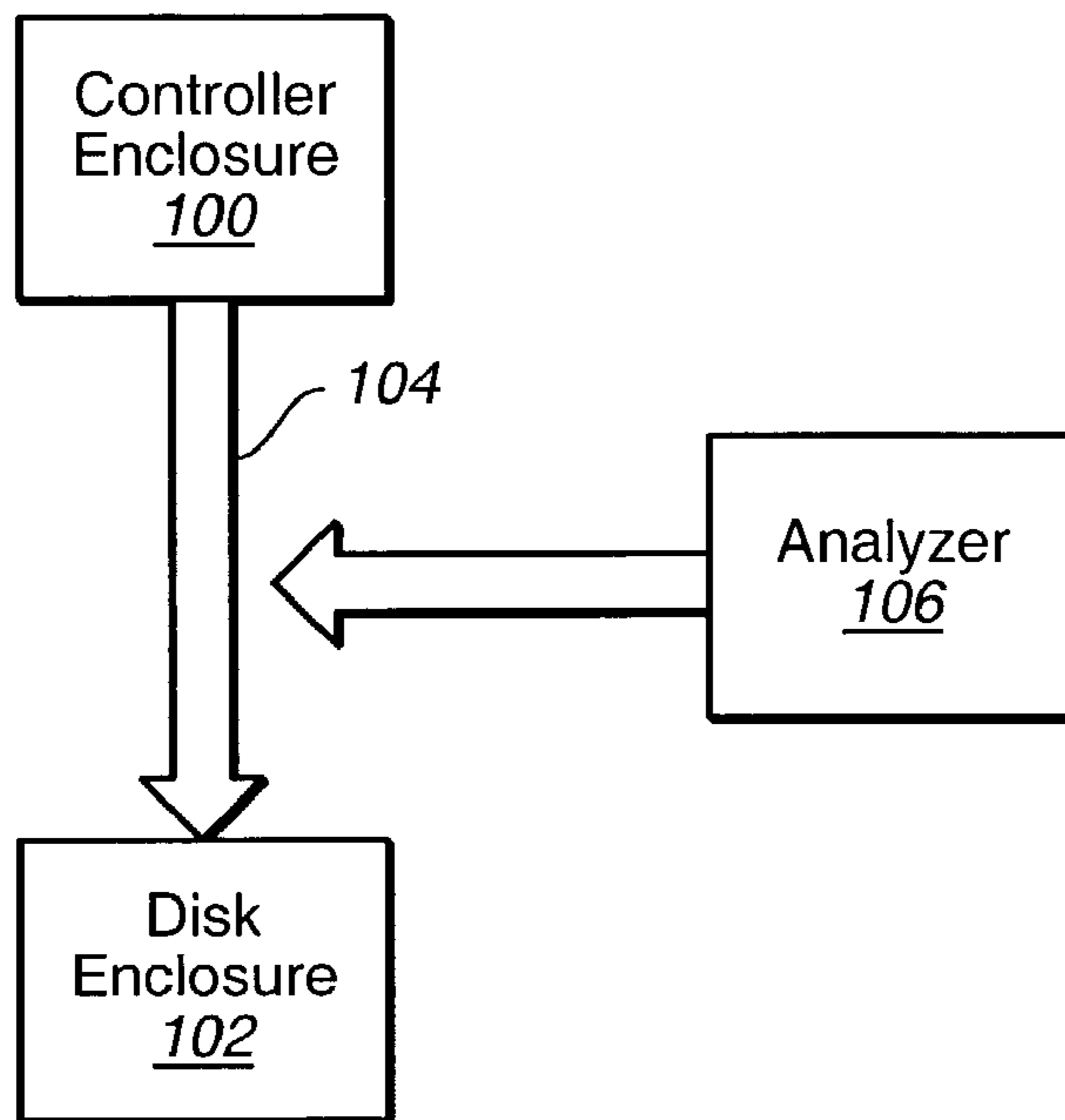
(74) *Attorney, Agent, or Firm*—Yee & Associates, P.C.

(57) **ABSTRACT**

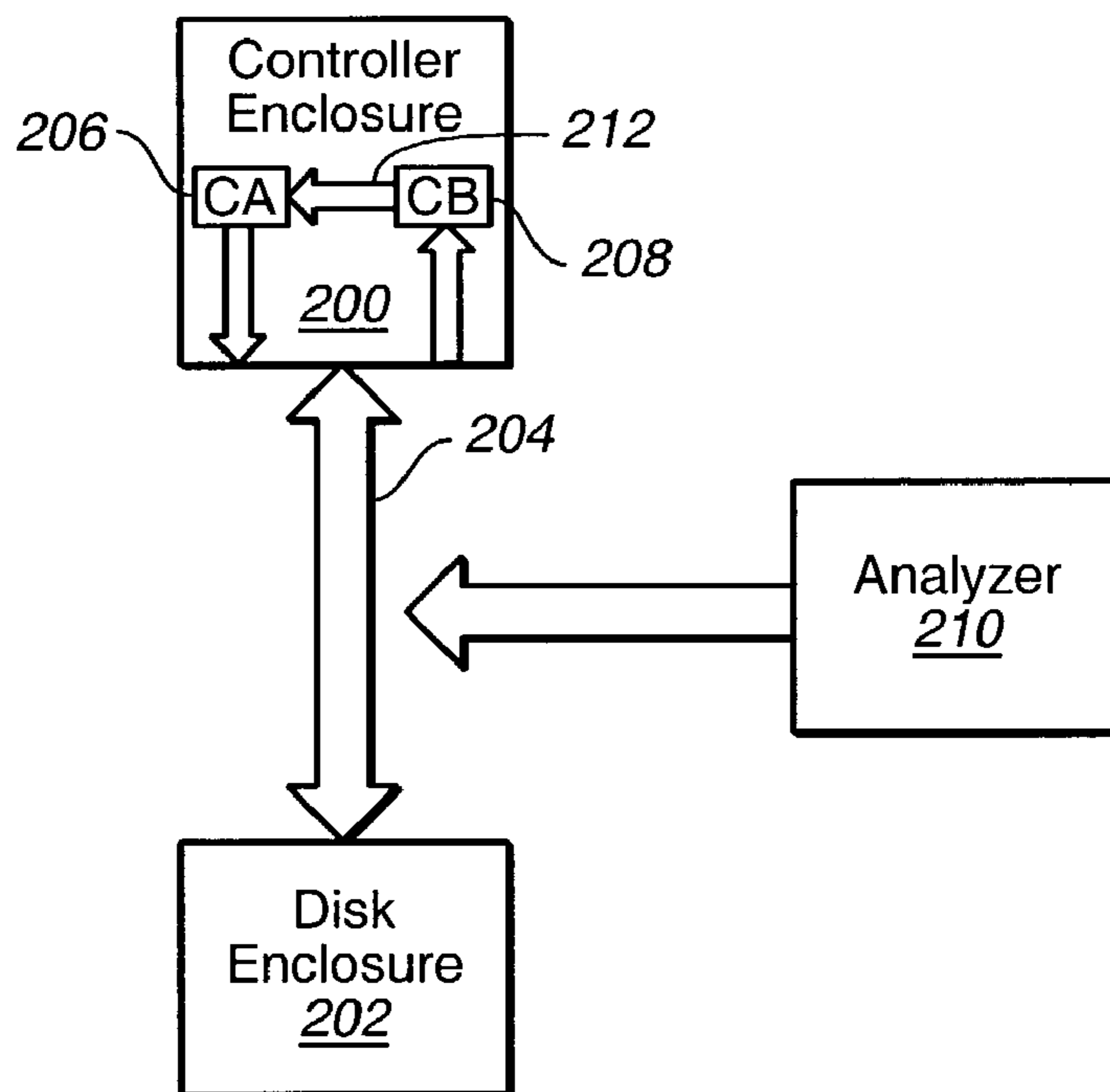
A method, apparatus, and computer instructions for a storage subsystem. This subsystem includes controller devices, storage devices, and a communications network. The communications network connects the controller devices and the storage devices. The communications network also includes a set of diagnostic outputs. The set of diagnostic outputs is configured to output data sent between two devices from the controller devices and the storage devices for monitoring.

**16 Claims, 5 Drawing Sheets**

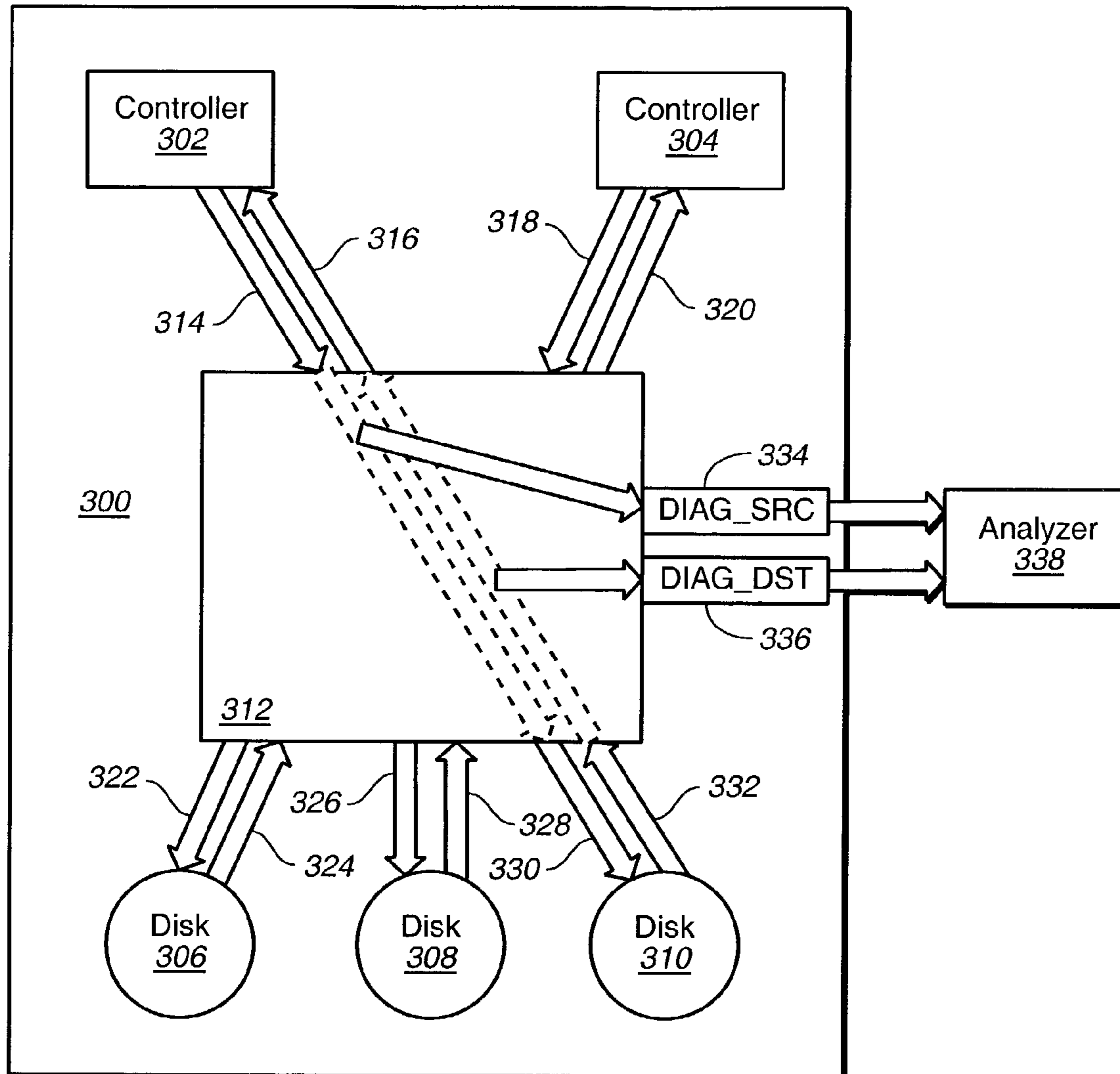




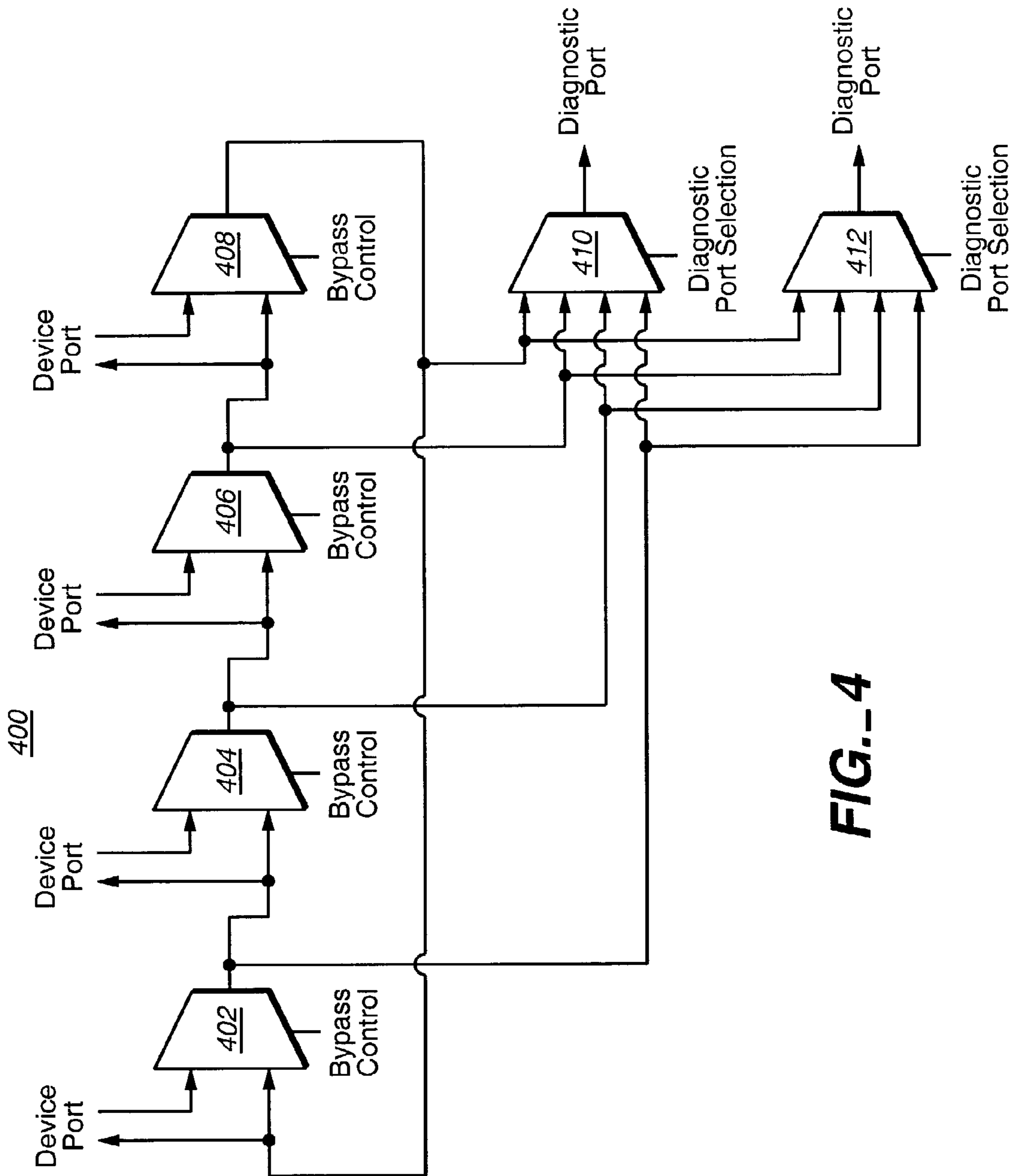
**FIG. 1**  
**(PRIOR ART)**



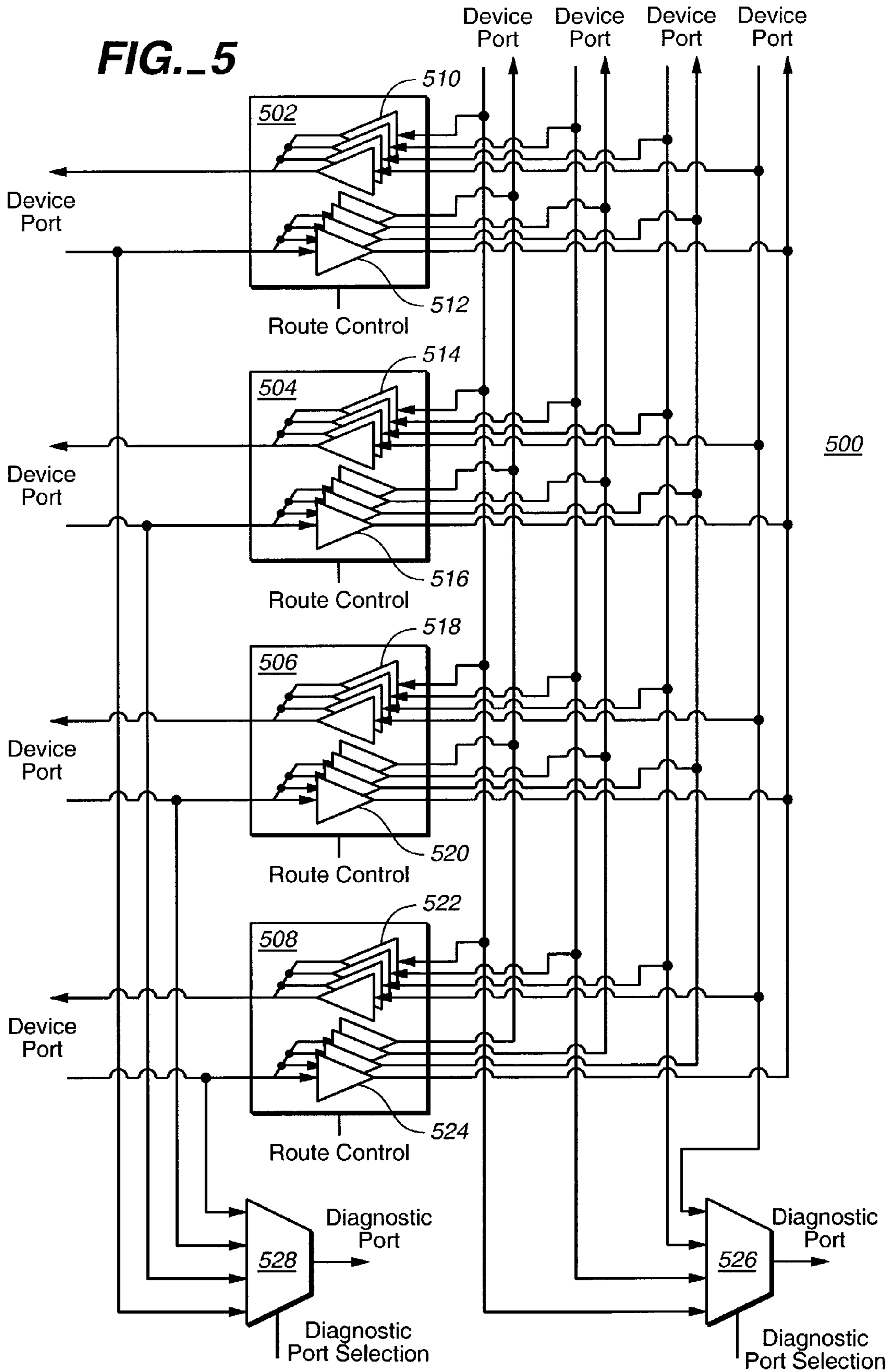
**FIG. 2**  
**(PRIOR ART)**

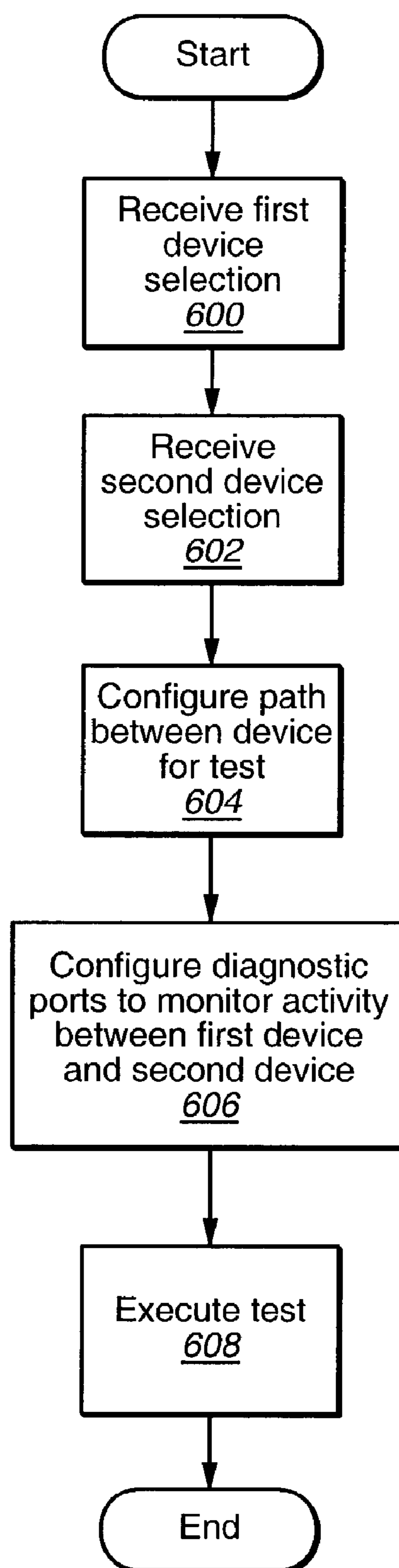


**FIG. 3**



**FIG.-4**



**FIG. 6**

1

**METHOD AND APPARATUS FOR  
DEBUGGING PROTOCOL TRAFFIC  
BETWEEN DEVICES IN INTEGRATED  
SUBSYSTEMS**

BACKGROUND OF THE INVENTION

1. Technical Field

The present invention is directed generally toward a method and apparatus for processing data. Still more particularly, the present invention provides a method, apparatus, and computer implemented instructions for debugging and performing diagnostic testing in a data processing system.

2. Description of the Related Art

Fibre channel is a high-speed transport technology used to build storage area networks (SANs). Fibre channel is a channel-network hybrid, containing sufficient network features to provide needed connectivity, distance, and protocol multiplexing, as well as sufficient traditional channel features to retain simplicity, repeatable performance, and guaranteed delivery. Although Fibre channel can be used as a general-purpose network carrying Asynchronous Transfer Mode (ATM), Internet Protocol (IP) and other protocols, this architecture has been primarily used for transporting Small Computer System Interface (SCSI) traffic from servers to disk arrays. The Fibre Channel Protocol (FCP) serializes SCSI commands into Fibre channel frames. IP, however, is used for in-band Simple Network Management Protocol (SNMP) network management. Fibre channel not only supports single-mode and multimode fiber connections, but coaxial cable and twisted pair as well.

Fibre channel can be configured point-to-point, via a switched topology or in an arbitrated loop (FC-AL) with or without a hub. Fibre channel provides both connection-oriented and connectionless services. With respect to the use of fibre channel in storage subsystem environments, such as those contained in disk arrays, debugging of these systems is currently formed by connecting an analyzer to extract traffic moving between devices for examination. For example, FIG. 1 is a diagram of a known controller/storage device system employing fibre channel. Controller enclosure 100 is connected to disk enclosure 102 through fibre channel connection 104. In this example, disk enclosure 102 is external to controller enclosure 100. Analyzer 106 is connected to fibre channel connection 104 to obtain data being transferred between these two devices for use in debugging problems manifesting themselves in communications between controller enclosure 100 and disk enclosure 102. In this simple example, obtaining data for analysis is straightforward and easy to set up.

In other cases, such an analysis is more difficult to perform. It is common for multiple devices to communicate with each other and be located in the same physical box or enclosure. For example, in FIG. 2, a diagram illustrating a known storage subsystem is illustrated. Controller enclosure 200 is connected to disk enclosure 202 by fibre channel connection 204. In this example, controller enclosure 200 contains two controllers, controller 206 and controller 208. These two controllers may initiate commands to disk enclosure 202 as well as send commands to each other. Analyzer 210 may be attached to fibre channel connection 204, which allows gathering data on traffic sent between controllers 206 and 208 and disk enclosure 202. In this case, however, data or traffic sent between controller 206 and controller 208 cannot be obtained by analyzer 210 because no physical attachment is present to connect analyzer 210 to link 212, which exists internally within controller enclosure 200. In this situation, trace logs

2

recorded by each of the controllers may be examined. This solution, however, requires gathering large amounts of data, which is often memory intensive. Further, no guarantee is present that any level of logging will provide sufficient detail to solve the problem. This situation is especially true when a hardware device, such as a protocol chip, is faulty and causes a bug. As a result, debugging is often an iterative and time-consuming process.

Therefore, it would be advantageous to have an improved method, apparatus, and computer instructions for analyzing and debugging problems in a storage subsystem.

SUMMARY OF THE INVENTION

The present invention provides a method, apparatus, and computer instructions for a storage subsystem. This subsystem includes controller devices, storage devices, and a communications network. The communications network connects the controller devices and the storage devices. The communications network also includes a set of diagnostic outputs. The set of diagnostic outputs is configured to output data sent between two devices from the controller devices and the storage devices for monitoring.

BRIEF DESCRIPTION OF THE DRAWINGS

The novel features believed characteristic of the invention are set forth in the appended claims. The invention itself however, as well as a preferred mode of use, further objects and advantages thereof, will best be understood by reference to the following detailed description of an illustrative embodiment when read in conjunction with the accompanying drawings, wherein:

FIG. 1 is a diagram of a known controller/storage device system employing fibre channel;

FIG. 2 is a diagram illustrating a known storage subsystem;

FIG. 3 is a diagram illustrating a storage subsystem in accordance with a preferred embodiment of the present invention;

FIG. 4 is a diagram illustrating a communications network using a loop topology in accordance with a preferred embodiment of the present invention;

FIG. 5 is a diagram illustrating a communications network using a switched network topology in accordance with a preferred embodiment of the present invention; and

FIG. 6 is a flowchart of a process used for gathering and analyzing data in accordance with a preferred embodiment of the present invention.

DETAILED DESCRIPTION

With reference now to the figures and in particular with reference to FIG. 3, a diagram illustrating a storage subsystem is depicted in accordance with a preferred embodiment of the present invention. In this example, storage subsystem 300 includes controller 302 and controller 304, which are used to control disk 306, disk 308, and disk 310. These devices are all connected through communications network 312. This communications network may take various forms, such as, for example, an arbitrated loop or a switched network. Storage subsystem 300 is implemented using fibre channel connections. The connections to devices are through device ports. In these examples, one device port is for receiving traffic or data, while the other device port is for transmitting traffic or data. Controller 302 is connected to communications network 312 through connections 314 and 316. Controller 304 is connected to communications network 312

through connections 318 and 320. Disk 306 is connected to communications network 312 through connections 322 and 324; disk 308 is connected to communications network 312 through connections 326 and 328; and disk 310 is connected to communications network 312 through connections 330 and 332.

Communications network 312 also includes diagnostic ports 334 and 336, which provide a port to connect analyzer 338 to storage subsystem 300 to gather data for debugging and diagnostic purposes. Diagnostic port 334 is used to output source data, while diagnostic port 336 is used to output destination data in this illustration. As shown, these diagnostic ports are configured to allow the gathering of data sent between controller 302 and disk 310. Controller 302 acts as an initiator with disk 310 being the target in this example. These diagnostic ports are not static, but may be configurable to provide for the gathering of data or traffic sent between any of the devices connected by communications network 312. For example, the same diagnostic ports, diagnostic port 334 and diagnostic port 336, may be configured to output data sent between other devices, such as controller 302 and controller 304.

Through this type of architecture in which a switching structure connects all of the devices to each other, a single set of diagnostic ports or output ports may be used to gather data between any combination of devices in the subsystem. This architecture allows transmissions or data between any initiator and any target to gather adequate data in monitoring system traffic. As a result, diagnostic port 334 and diagnostic port 336 may be a data outlet for any device connected to communications network 312. This flexibility allows for analysis of traffic between all devices without requiring a separate port for each transmission line as currently required in presently used architectures, such as the system illustrated in FIG. 2.

Turning now to FIG. 4, a diagram illustrating a communications network using a loop topology is depicted in accordance with a preferred embodiment of the present invention. Communications network 400 uses a loop network topology. Communications network 400 is an example of a particular implementation of communications network 312 in FIG. 3. Traffic between different devices is controlled through the use of multiplexers 402, 404, 406, and 408. These multiplexers are configured in a serial loop configuration. In these examples, each multiplexer contains two inputs and one output. The output of multiplexer 402 is connected to the input of multiplexer 404, the output of multiplexer 404 is connected to the input of multiplexer 406, and the output of multiplexer 406 is connected to the input of multiplexer 408. The output of multiplexer 408 loops back and is connected to the input of multiplexer 402 to form the loop configuration. The other input of each multiplexer is connected to a device port in a device, such as, for example, controller 302 or disk 306 in FIG. 3. Further, the output of each multiplexer is connected to a device port. The input selected to be output through the output of a multiplexer is selected through a bypass control signal.

In addition, communications network 400 also includes multiplexer 410 and multiplexer 412. The output of these multiplexers form a pair of diagnostic ports, such as diagnostic port 334 and 336 in FIG. 3. The inputs of each of these diagnostic ports is connected to an output of each of the multiplexers. The input selected for output to a diagnostic port is selected through a diagnostic port selection signal. Through these selection signals, the traffic from any combination of devices connected to communications network 400 may be output to diagnostic ports for analysis. This architec-

ture does not require a port or output for each path or connection between different devices within the subsystem.

With reference now to FIG. 5, a diagram illustrating a communications network using a switched network topology is depicted in accordance with a preferred embodiment of the present invention. Communications network 500 is an example of a particular implementation of communications network 312 in FIG. 3. Communications network 500 contains switching units 502, 504, 506, and 508. Each of these switching units is connected to a set of device ports in the communications network. Each switching unit contains a set of input buffers and a set of output buffers. As illustrated, switching unit 502 contains input buffers 510 and output buffers 512, switching unit 504 contains input buffers 514 and output buffers 516; switching unit 506 contains input buffers 518 and output buffers 520; and switching unit 508 contains input buffers 522 and output buffers 524. The input buffers are connected to the set of device ports and converge to a single output for connection to a device port. The output buffers are connected to the set of device ports with the inputs to these buffers being connected to a single device port. A control mechanism is used to select a particular buffer for output or input. In this example, four device ports are present. The particular device ports selected are controlled by a route control signal.

Communications network 500 also includes multiplexer 526 and multiplexer 528, which provide an output for connection to a diagnostic port, such as diagnostic port 334 and 336 in FIG. 3. The inputs of multiplexer 526 are connected to the input having a path through the output buffers, while multiplexer 528 has its inputs connected to the inputs for the input buffers in each switching unit. The particular input selected for output through a multiplexer is controlled through a diagnostic port selection signal. In this manner, communications network 500 may provide an ability to output data for diagnostic and debugging uses through a single set of diagnostic ports. Having a diagnostic port for each possible path or route between different devices is avoided through this configurable diagnostic port mechanism implemented within communications network 500.

With reference now to FIG. 6, a flowchart of a process used for gathering and analyzing data is depicted in accordance with a preferred embodiment of the present invention.

The process illustrated in FIG. 6 may be implemented in a storage subsystem, such as storage subsystem 300 in FIG. 3. In particular, these steps may be implemented in the form of computer instructions or firmware to generate the appropriate control signals for configured outputs for a set of diagnostic ports, such as diagnostic port 334 and diagnostic port 336 in FIG. 3.

The process begins by receiving the first device selection (step 600). A second device selection is received (step 602). These device selections may be for any initiator and target in a storage subsystem. A path is configured between the device for testing (step 604). This path may be configured by sending or generating the control signals for the communications network to select or connect the devices to each other in a path through the communications network. Diagnostic ports are configured to monitor activity between the first device and the second device (step 606). The configuration is achieved through sending or generating the appropriate diagnostic port selection signals such that the output in the diagnostic ports corresponds to the path between the selected devices. A test is executed (step 608). This test includes gathering data sent between the selected devices and analyzing the data. An example of an analyzer, which may be used to obtain and



## 5

analyze data is a Finisar GTX Fibre Channel Analyzer, which is available from Finisar Corporation.

Thus, the present invention provides an improved method, apparatus, and computer instructions for monitoring the transfer of data between different devices. Specifically, commands sent between devices may be gathered for diagnostic and debugging purposes. The mechanism of the present invention avoids having to require a separate port for every path or transmission between different devices in a storage subsystem. A communications network implementing a configurable port system is included in the storage subsystem to interconnect the different devices. Thus, each time a particular initiator and target is selected for monitoring, the ports may be configured to output data sent between those two devices.

The description of the preferred embodiment of the present invention has been presented for purposes of illustration and description, but is not intended to be exhaustive or limited to the invention in the form disclosed. Many modifications and variations will be apparent to those of ordinary skill in the art. For example, the subsystems illustrated are storage subsystems. Of course, the mechanism of the present invention may be applied to other types of subsystems other than storage subsystems. The mechanism of the present invention may be incorporated into any node based network protocol, including, for example, computer motherboard architectures such as Infiniband or PCI Express to provide a diagnostic analyzer access to the bus. The embodiment was chosen and described in order to best explain the principles of the invention the practical application to enable others of ordinary skill in the art to understand the invention for various embodiments with various modifications as are suited to the particular use contemplated.

What is claimed is:

1. A storage subsystem for gathering and analyzing data sent between devices, the storage subsystem comprising:

a plurality of controller devices;

a plurality of storage devices;

a communications network connecting the plurality of controller devices and the plurality of storage devices, wherein a pair of controller devices, a pair of storage devices, and a combination of a controller device and a storage device each form a selected pair of devices;

a set of diagnostic outputs defined in the communications network;

a controller connected to the communications network, wherein the controller selectively configures the set of diagnostic outputs to receive data sent between devices in any selected pair of devices for the gathering and analyzing of the data sent between the devices in the selected pair of devices, wherein data from a selected device in the pair of selected devices is directed to a corresponding diagnostic output and data from a remaining selected device in the pair of selected devices is directed to another corresponding diagnostic output in the set of diagnostic outputs; and

an analyzer capable of analyzing the data from the set of diagnostic outputs.

2. The storage subsystem of claim 1, wherein the communications network has a loop network topology.

3. The storage subsystem of claim 1, wherein the communications network has a switched network topology.

4. The storage subsystem of claim 1, wherein the plurality of storage devices includes at least one of a hard disk drive, tape drive, and an optical drive.

5. The storage subsystem of claim 1, wherein the set of diagnostic outputs is a pair of diagnostic outputs.

## 6

6. The storage subsystem of claim 1, wherein the set of diagnostic outputs output data sent between a first controller device and a second controller device within the plurality of controller devices.

7. The storage subsystem of claim 1, wherein the set of diagnostic outputs output data sent between a controller device within the plurality of controller devices and a storage device within the plurality of storage devices.

8. The storage subsystem of claim 1, wherein the set of diagnostic outputs comprises:

a first diagnostic port to output source data of one device in the selected pair of devices; and

a second diagnostic port to output destination data of a remaining device in the selected pair of devices.

9. The storage subsystem of claim 1, wherein the set of diagnostic outputs is selectively configurable to receive and output data from between any two devices of the storage subsystem.

10. The storage subsystem of claim 9, wherein the set of diagnostic outputs is further configurable to analyze data traffic between any two devices absent a need for a separate port per transmission line.

11. The storage subsystem of claim 1, wherein the communications network further comprises:

a plurality of multiplexers, configured in a loop configuration to control data traffic in the storage subsystem, wherein each multiplexer has two inputs and one output, the output of one multiplexer connected to the input of another multiplexer to form the loop configuration in which each multiplexer is further connected to each diagnostic output of the set of diagnostic outputs; and a bypass control signal to select an input to be output through an output of a multiplexer in the plurality of multiplexers.

12. The storage subsystem of claim 1, wherein the communications network further comprises:

a plurality of switching units, wherein each switching unit in the plurality of switching units containing a set of input buffers connected to a set of device ports to converge to a single output for connection to a device port and a set of output buffers connected to the set of device ports having input to the set of output buffers connected to a single device port; and

a control mechanism to select a buffer for one of input and output.

13. The storage subsystem of claim 12, wherein the control mechanism is a route signal.

14. The storage subsystem of claim 12, wherein the communications network further comprises:

a first multiplexer having input connection through the set of output buffers;

a second multiplexer having input connection through the set of input buffers; and

a diagnostic port selection signal controlling input selection for output through each multiplexer.

15. A storage subsystem for gathering and analyzing data sent between devices, the storage subsystem comprising:

a plurality of controller devices;

a plurality of storage devices;

a plurality of switching devices;

a plurality of multiplexer devices; and

a communications network connecting the plurality of controller devices and the plurality of storage devices using

7

the plurality of switching devices and plurality of multiplexer devices, wherein a pair of controller devices, a pair of storage devices, and a combination of a controller device and a storage device each form a pair of devices; and

a set of diagnostic outputs defined in the communications network;

a controller connected to the communications network, wherein the controller selectively configures the set of diagnostic outputs to receive data sent between devices in any selected pair of devices for the gathering and analysis of data sent between the devices in the selected pair of devices, wherein data from a selected device in the pair of selected devices is directed to a corresponding diagnostic output and data from a remaining selected

8

device in the pair of selected devices is directed to another corresponding diagnostic output in the set of diagnostic outputs; and

an analyzer capable of analyzing the data from the set of diagnostic outputs, wherein output of the diagnostic outputs corresponds to a path through the selected devices.

**16.** The storage subsystem of claim **15** wherein the communications network selectively configures the set of diagnostic ports and further comprises:

a control mechanism in the form of one of a route signal for selecting a device port and a diagnostic port selection signal for selecting a particular input for output through a respective multiplexer in the plurality of multiplexer devices.

\* \* \* \* \*